

The Biological Bases of Syntax-Semantics Interface in Natural Languages: Cognitive Modeling and Empirical Evidence.

Malaia, E., Wilbur, R.B.

SLHS Department and Linguistics Program, Purdue University
500 Oval Drive, West Lafayette, IN 47907
emalaya@purdue.edu, wilbur@purdue.edu

Abstract

This paper reviews recent research on representation of events in human languages. We consider empirical evidence for event-structural analysis of language comprehension and production in spoken and signed languages, as well as possible biological bases for it. Finally, theoretical linguistic models, models of language processing, and cognitive architectures, which account for such event-structural basis of syntax-semantics (and, possibly, phonology interface in ASL) in human languages are discussed.

Representation of events in human languages: linguistic universals meet language processing.

The idea that human languages parse and formulate observable events in a logically restricted fashion is fairly old, dating back to Vendler's (1967) Aktionsart predicate classes (more recently developed by van Lambalgen and Hamm, 2005). Recent work by Van Valin (2007) claims that the most pervasive components of real-world events have made their way into the morphology of most of the world's languages (albeit in different form), qualifying them for the status of linguistic universals. For example, he notes that durative, dynamic, causative, and inchoative alternations surface as morphological constants in various language families; while the granularity of the system remains to be (empirically) determined, the pervasiveness of it in the linguistic data is striking.

Linguistic theory of verb types has long observed universal correspondences between verbal meanings and syntactic behaviors, including adverbial modification (Tenny 2000), aspectual coercion (Smith, 1991), and argument structure alternations (Levin, 1993, Ramchand, 2008, etc.). Vendler (1967) proposed a system of four basic syntactically relevant semantic types of predicates: atelic States and Activities, and telic Achievements and Accomplishments. The telic/atelic distinction is most clearly analyzed in terms of the internal structure of events. 'Telic' is understood as the property of events (linguistic predicates) containing a conceptual (semantic) endpoint. In contrast, 'atelic' events do not contain such a point and have the potential to continue indefinitely. Atelic events are homogenous, in that they may be divided into identical intervals, each of which is an instance of the event itself, i.e. 'walking' as an

instance of 'walking'. Telic events are composed of at least two subevents, one of which is the final state, and are therefore heterogeneous (cannot be divided into identical intervals). The model was further developed by Pustejovsky (2001), with the primary distinction between static subevent type S(tate) and dynamic subevent type P(rocess). Vendlerian transitions were modeled as combinations of non-identical subevents: either $S \rightarrow S$ (Achievements) or $P \rightarrow S$ (Accomplishments). Ramchand (2008) further developed a syntax-semantics interface model of event and argument structure including three possible event phases: initiation, process, and resultant state (with corresponding argument realization), and elaborated a factorial typology of event-argument predicate constructs based on formal minimalist syntax, overtly incorporating semantic notions of durativity, resultant state/telicity, agentivity and causation into the interface.

Linguistic theories beyond minimalism also converge on a similar set of semantic distinctions affecting syntax in the world's languages. For example, Dixon (2000) lists state/action, control, volition, affectedness and intention among semantic primitives involved in contrastive pairs of causal constructions across the world's languages; and Pinker (1989), in elaboration of the Grammatically Relevant Semantic Subsystem Hypothesis (GRSSH), distinguishes abstract semantic features, which affect grammar (e.g. punctual vs. durative, causation), and idiosyncratic ones (such as manner), which do not.

While complete analysis of all known languages of the world with respect to their event structure is not currently feasible, additional evidence for the fundamental character of a predicate's event structure comes from psycholinguistic and neurolinguistic research; an increasing body of research is supplying empirical evidence that such concepts as telicity affect the way syntactic structure is processed in human languages.

For example, O'Bryan (2003) showed independent effects of telicity and transitivity on response times in a word maze experiment with reduced relative clauses. The subjects were asked to complete grammatical sentences containing Object reduced relative clauses, such as "The actress awakened by the writer left in a hurry". There was a significant reaction time advantage on the preposition "by" for sentences with telic verbs in the relative clause, as

compared to those with atelic verbs; there was also an independent advantage for the second argument in sentences with obligatorily transitive verbs (both telic and atelic). The experiment thus demonstrated independent effects of telicity and transitivity on human language processing ability. A similar study by Friedmann, Taranto, Shapiro, and Swinney (in press) compared sentences with unergative (intransitive atelic) and unaccusative (intransitive telic) English verbs using a cross-modal priming technique, and found a priming effect for non-alternating unaccusatives (obligatorily intransitive telic verbs), but not for unergatives (obligatorily intransitive atelics).

Additional evidence of telicity affecting online sentence processing in spoken English was provided in an EEG study by Malaia, Wilbur, and Weber-Fox (2008), which investigated the effects of verbal telicity on the ease of syntactic re-analysis of Object reduced relative clauses. The ERP data demonstrated that subjects' recovery from garden-path effects required fewer processing resources when telic verbs were used in reduced relative clauses. The same frame alternations in atelic verbs were more difficult to process, and elicited ERP waveforms typically associated with increased difficulty of early syntactic processing and thematic role integration. Because the study controlled for other factors which could affect processing, including argument animacy, frequency of verb occurrence, and transitivity, the results were clearly due to the linguistic telicity influencing sentence processing at the syntax-semantics interface.

Theoretical models which account for such empirical data range from construction grammar (Kemmerer 2006, Goldberg 1995), claiming regular correspondence between semantics and syntax in verbal constructions, to a fully developed event-structure based syntax-semantics interface (Ramchand 2008), which attempts to account for the ways regular constructions and cross-linguistically observable morphology could have developed, using the minimalist model of syntax. Psycholinguistic models of processing grammatically relevant semantic information, on the other hand, are still under development (see Kemmerer, 2006, for examples). Even less is known about emergence of grammatically relevant semantic features in the lexicon. We suggest however, that signed languages, which are more closely tied to the visual modality - simultaneously the one of perception - might provide the missing link to the puzzle.

Overt representation of event structure in signed languages.

While spoken languages cross-linguistically demonstrate consistent utilization of event structure in morphology and the syntax-semantics interface (Folli & Harley, 2006, for review), an even more compelling pattern emerges in signed languages. Because of the visual modality of signing matches a major human perceptual interface, many

signs are not as arbitrary in their construction/form as spoken words are. While signs are not literally 'iconic' (otherwise different sign languages would adopt the same exact signs for the same concepts, something classically demonstrated to be untrue in Klima & Bellugi, 1979), it is much more likely that the fundamental features based on human event perception might be found in the grammar and phonology of a signed language, as argued in Wilbur (2003, in press).

From the standpoint of theoretical linguistics, the event-based analysis of predicate semantics and syntax has been supported in general for ASL (Wilbur, 2003; Rathmann, 2005). Semantics of event type has been shown to have a direct effect not only on the number of arguments, but on the available morphological modifications as well: Brentari (1998) notes that delayed completive aspect marking only applies to telic stems; and Wilbur (2003) argues that the phonological structure of predicate signs in ASL is compositional, and that the components are grammaticalized from universally available physics of motion and geometry of space. Schalber (2006) observes this mapping in Austrian Sign Language, not only in the form of manual signs but also in the behavior of associated mouth movements.

The direct mapping of perception-based event parsing to ASL predicate phonology/semantics/syntax interface has been thus formulated as Event Visibility Hypothesis in Wilbur (2003): 'movement which stops at points (p) in space maps semantically to the final State of telic events (en) and its individual argument semantic variable (x)'. The EVH predicts that telic predicates are expected to be found with noticeably more rapid movement to a stop, marking the final state, as compared to atelic predicates.

The kinematics of production proposed by hypothesis was tested in a motion capture experiment in Wilbur & Malaia (2008). In this experiment, a group of telic and atelic signs were randomized, and presented to a native ASL signer, who produced them in isolation, in the carrier phrase 'SIGN X AGAIN', sentence-medially 'SHE X TODAY', and sentence-finally 'TODAY SHE X'. Both the video and the vectors for the 3-dimensional location, velocity and acceleration of the right wrist marker were imported into video annotation software and aligned using the audio marker. The video was annotated marking the beginning of each predicate (once the dominant hand assumed appropriate handshape) to the end of each movement (point of contact, or maximal displacement). The data from the right wrist marker indicated that in all environments, mean deceleration of telic signs was 1.5 to 2 times steeper than that of atelic signs, with a significant difference between telic and atelic signs ($p < 0.05$). The experiment demonstrated that there are, indeed, production differences in ASL predicate signs reflecting syntactico-semantic telicity feature: ASL takes advantage of available kinematic distinctions to provide cues to the viewer regarding the event structure of the predicate. From the linguistic standpoint, such overt difference in sign

production mapping onto event-structural representation of syntax-semantics interface has implications for modeling the syntax-semantics interface in both signed and spoken languages. However, availability of kinematic data on ASL suggested additional lines of inquiry. For example, are telic and atelic predicates in ASL processed differently on the neuroanatomical level? And what perceptual features of signs might facilitate such differential processing?

Neuroanatomical substrates underlying event structure processing.

The question of neuroanatomical correlates for differential comprehension of telic and atelic predicates in American Signed Language was addressed in a pilot fMRI study (Malaia, Wilbur, Talavage, in press), investigating neurological subsystems underlying event structure processing. Previous studies on spoken languages have demonstrated that left inferior frontal gyrus (IFG) is activated during both syntactic and semantic processing, leading to the hypothesis that it may be the locus of integration of syntactic and semantic information during sentence comprehension (Hagoort 2005). Emmorey's (2006) review of Broca's area in sign language research also notes that ASL production studies consistently observe left hemispheric activation of Broca's area (cf. Corina et al. 1999, Horowitz et al. 2003, Emmorey 2002, 2003, 2004). ASL comprehension studies, on the other hand, have mostly found bilateral activation in Broca's area (Levänen et al, 2001, Neville et al., 1998, etc.), which is also typical for audiovisual stimuli comprehension in spoken languages (cf. Capek et al., 2004). Recruitment of left hemisphere Broca's region for sign language production does not appear to depend on signer handedness (Corina et al., 2003), but classifier constructions in ASL (which do not require lexical retrieval per se) do not engage Broca's area significantly (Emmorey and Herzig, 2003); thus, spoken and signed languages appear to be highly correlated in the use of Broca's area in the left hemisphere for integration of structural and lexical linguistic information. Emmorey (2007) has also demonstrated IFG activation during ASL processing, as compared to pantomime. Given the theoretical accounts of the role telicity plays at the syntax-semantic interface, our study predicted differentiated activation patterns in IFG (BA 44/45) for telic and atelic ASL signs, when contrasted using a subtraction paradigm.

The subjects in the study were presented with visual stimuli consisting of telic and atelic ASL signs in a block paradigm, with simple gestures as a baseline condition¹. Semantic processing of the stimuli was ensured by the task, during which the subjects had to indicate whether the

action denoted by the predicate was more likely to occur inside a house, or outside, by pressing a button. Fixed effects analysis demonstrated activation clusters in Broca's area (BA 44/45, centered at MNI [-52 24 0], cluster size 15 voxels), and angular gyrus (BA39, cluster size 5 voxels, centered at MNI [46 33 13]); both clusters survive FDR thresholding at $p < 0.05$. Although more data is needed to determine the generalizability of the findings, the preliminary result supports the hypothesis that semantic information encoded in event structure of the ASL predicates is relevant for integration of semantic and syntactic information, and that parametric differences in production of the signs in the motion capture experiment are available for perception and used for integrative online semantic processing².

These findings are consistent with other neuroimaging research, in both signed and spoken languages, as noted above. Additionally, the evidence that the predicates which differ in telicity feature appear to differentially engage Broca's area highlights the theoretical relevance of event structure modeling for language processing. The combined results of the motion capture and neuroimaging experiments in ASL confirm the prediction that world-view information is systematically recruited by American Sign Language for sign production, and is used by native signers in perception of predicate signs; our results also point to early interaction of syntax and semantics in human languages, consistent with lexicalist parsing models (cf. Jackendoff, 2002), and especially the Unification Model (Vosse and Kempen, 2000). As Hagoort (2006) points out in elaboration of the Unification Model of language processing, "Combinatoriality... holds equally for syntactic, semantic, and phonological levels of analyses". Thus, for signed language, modeling of event structure appears to unite all three levels of linguistic representation, and present a case for a linguistic universal that is mandatory to model in cognitive architectures.

What follows from empirical evidence in signed and spoken languages is that not only semantics, but also syntax of human languages cross-modally is grounded in what can be construed as biological perception. In other words, the complexity of the interaction between semantics and syntax is not limited to consistent occurrence of certain structures in a specific language, as claimed by constructionist approaches (Goldberg 1995, etc.), but rather operates through the complexity of linguistic structures. We suggest that events in the real world are perceived, conceptualized and verbalized in a way which takes advantage of syntax-semantics interface with the built-in account of real-world events – which is a ready-made framework for a computational system (cf. van Lambalgen, 2005).

¹ The simple gestures consisted of non-intentional (non-communicative) slow movements of the upper arms being lowered from a t-position (straight out) to the sides of the body, and back up.

² This pilot study did not specifically address the question of perceptual features relevant for processing, which would have required a different experimental design and control conditions.

References

- Brentari, D. 1998. A prosodic model of sign language phonology, Cambridge, MA: MIT Press.
- Capek, C.M., Bavelier, D., Corina, D., Newman, A.J., Jezzard, P., & Neville, H.J., 2004. The cortical organization of audio-visual sentence comprehension: An fMRI study at 4 Tesla. *Cognitive Brain Research*, 20, 111-119.
- Corina, D.P., San Jose-Robertson, L., Guillemin, A., High, J., & Braun, A. 2003. Language lateralization in a bimanual language. *Journal of Cognitive Neuroscience*, 15, 718-730.
- Corina, D.P., McBurney, S.L., Dodrill, C., Hinshaw, K., Brinkley, J., & Ojemann, G. 1999. Functional roles of Broca's area and supramarginal gyrus: Evidence from cortical stimulation mapping in a deaf signer. *NeuroImage*, 10, 570-581.
- Dixon, R.M.W. 2000. A typology of causatives: form, syntax, and meaning. In Dixon, R.M.W. and Aikhenvald, A.Y. (eds). *Changing valency*. Cambridge, UK: Cambridge University Press, pp. 30-83.
- Emmorey, K. 2002. *Language, cognition, and the brain: Insights from sign language research*. Mahwah, NJ: Lawrence Erlbaum and Associates.
- Emmorey, K. (ed.) 2003. *Perspectives on classifier constructions in sign languages*. Mahwah, NJ: Lawrence Erlbaum and Associates.
- Emmorey, K., Grabowki, T., 2004. Neural organization for sign versus speech production. *Journal of Cognitive Neuroscience Supplement*, F69, 205.
- Emmorey, K. 2006. Role of Broca's area in sign language. In *Broca's Region*, Grodzinsky, Y. & Amunts, K. eds., Oxford University Press
- Emmorey, K., Xu, J., Gannon, P., Goldin-Meadow, S., & Braun, A. 2007. Perception of pantomime, American Sign Language verbs, and nonsense signs by deaf signers and hearing non-signers. *Cognitive Neuroscience Society*, May, New York.
- Emmorey, K., Herzig, M., 2003. Categorical versus gradient properties of classifier constructions in ASL. In K. Emmorey (Ed), *Perspectives on classifier constructions in sign languages*. Mahwah, NJ: Lawrence Erlbaum and Associates.
- Friedmann, N., Shapiro, L.P., Taranto, G., and Swinney, D. (in press). The leaf fell (the leaf): the on-line processing of unaccusatives. *Linguistic Inquiry*.
- Goldberg, A. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, IL: University of Chicago Press.
- Hagoort, P. 2006. On Broca, brain, and binding. In *Broca's Region*, Grodzinsky, Y. & Amunts, K. eds., Oxford University Press
- Horowitz, B., Amunts, K., Bhattacharyya, R., Patkin, D., Jeffries, K., Zilles, K. 2003. Activation of Broca's area during the production of spoken and signed language: A combined cytoarchitectonic mapping and PET analysis. *Neuropsychologica*, 41, 1868-1876.
- Jackendoff, R., 2002. *Foundations of language: brain, meaning, grammar, evolution*. Oxford, UK: Oxford University Press.
- Kemmerer, D. 2006. Action verbs, argument structure constructions, and the mirror neuron system. In Arbib, M. (Ed). *Action to Language via the Mirror Neuron system*, Cambridge University Press, pp. 347-373.
- Levänen, S., Uutela, K., Salenius, S., & Hari, R. 2001. Cortical representation of sign language: Comparison of deaf signers and hearing non-signers. *Cerebral Cortex*, 11, 506-512.
- Levin, B. 1993. *English Verb classes and alternations*. The University of Chicago Press.
- Neville, H., Bavelier, D., Corina, D., Rauschecker, J., Kami, A., Lalwani, A., et al. 1998. Cerebral organization for language in deaf and hearing subjects: Biological constraints and effects of experience. *Proceedings of the National Academy of Science USA*, 95, 922-929.
- Malaia, E., Wilbur, R.B., Talavage, R., in press. Experimental evidence of event structure effects on American Sign Language predicate production and processing. *Chicago Linguistic Society* 44.
- O'Bryan, E. (2003). *Event Structure in Language Comprehension*. Unpublished manuscript, University of Arizona.
- Pinker, S. 1989. *Learnability and Cognition*. Cambridge, MA: MIT Press.
- Pustejovsky, J. 1991. The syntax of event structure. *Cognition*, 41(1-3): 47-81.
- Ramchand, G. 2008. *Verb Meaning and the Lexicon: A First Phase Syntax*, Cambridge: Cambridge University Press.
- Rathmann, C. 2005. *Event Structure in American Sign Language*, Ph.D. dissertation, University of Texas at Austin.
- Schalber, K. 2006. Event Visibility in Austrian Sign Language (ÖGS). *Sign Language & Linguistics*, 9(1/2), 207-231.
- Smith, C. S. 1991. *The Parameter of Aspect*. Dordrecht: Kluwer Academic Publishers.
- Tenny, C. 2000. Core events and adverbial modification. In Tenny & Pustejovsky (eds). *Events as grammatical objects*. Stanford, CA: CSLA Publications.
- van Lambalgen, M., Hamm, F. 2005. *The proper treatment of events*. Blackwell Publishing.
- Van Valin (2007) Some universals of verb semantics. In *Linguistic Universals*, Mairal, R., Gil, J. eds. Cambridge University Press.
- Vendler, Z. 1967. *Linguistics in Philosophy*, Cornell University Press, New York.
- Vosse, T., Kempen, G.A.M., 2000. Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and lexicalist grammar. *Cognition*, 75, 105-143.
- Wilbur, R. B. 2003. Representations of telicity in ASL. *Chicago Linguistic Society* 39 (1): 354-368.
- Wilbur, R., Malaia, E. (2008) Event Visibility Hypothesis: motion capture evidence for overt marking of telicity in ASL. *Linguistic Society of America meeting*, Chicago, IL.